

The HZSK Repository. Implementation, Features, and Use Cases of a Repository for Spoken Language Corpora

Daniel Jettka, Daniel Stein

D-Lib Magazine, ISSN 1082-9873, Vol. 20, n. 9/10, 2014

El artículo describe el proceso de concepción e implementación de un repositorio digital que se basa en las plataformas de software de Fedora, Islandora y Drupal. El tipo de dato principal que se va a organizar, mantener y distribuir con el repositorio habla un corpus lingüístico con un enfoque de multilingüismo, lo que significa que el contenido del repositorio puede ser de especial interés para los investigadores de lingüística, así como de estudios de traducción y socio-psicológicos. Nos concentramos en los requisitos técnicos que se tuvieron que cumplir en lo que respecta a nuestro status como centro oficial de la Common Language Resource Infrastructure Network (CLARIN), una iniciativa pan-europea en el contexto de las Humanidades Digitales. En particular, los requisitos incluyen la preparación de información para la descripción de los recursos que proporcionan una fuente de recolección de metadatos, registro de identificadores persistentes para el acceso a largo plazo y cita de recursos, implementación de un módulo de búsqueda de contenido federado, que se encarga de las restricciones de privacidad, y la posibilidad de proporcionar un único método de inicio de sesión vía Shibboleth. El Repositorio HZSK pasó satisfactoriamente dos procesos de evaluación: la evaluación del Data Seal of Approval, y un procedimiento interno de evaluación de CLARIN.

Resumen traducido de la propia publicación